13 Char_align: a program for aligning parallel texts at the character level
Kenneth Ward Church
June 1993 **Proceedings of the 31st annual meeting on Association for Computational Linguistics**
Full text available: pdf(639.85 KB)
Publisher Site

Additional Information: full citation, abstract, references, citings

There have been a number of recent papers on aligning parallel texts at the sentence level, e.g., Brown *et al* (1991), Gale and Church (to appear), Isabelle (1992), Kay and Rösenschein (to appear), Simard *et al* (1992), Warwick-Armstrong and Russell (1990). On clean inputs, such as the Canadian Hansards, these methods have been very successful (at least 96% correct by sentence). Unfortunately, if the input is noisy (due to OCR and/or unknown markup conventions), then these metho ...